DOCUMENT RESUME

ED 100 357 16 IR 001 455

AUTHOR Nickerson, Raymond S.; Stevens, Kenneth N.
TITLE Development of Computerized Speech-Training Aids for the Deaf. Report No. 2914. Final Report.
INSTITUTION Bolt, Beranek and Newman, Inc., Cambridge, Mass.
SPONS AGENCY Bureau of Education for the Handicapped (DHEW/OE), Washington, D.C. Media Services and Captioned Films Branch.
REPORT NO BBN-R-2914
PUB DATE 14 Sep 74
CONTRACT OEC-0-71-4670(615)
NOTE 50p.; For related documents see IR 001 449-454

EDRS PRICE MF-$0.75 HC-$3.15 PLUS POSTAGE
DESCRIPTORS *Computer Assisted Instruction; *Deaf Children; Deaf Education; Electromechanical Aids; Program Descriptions; Programed Materials; *Spectrograms; *Speech Handicaps; *Speech Instruction; Speech Skills; Speech Therapy; Teaching Techniques; Training Techniques; Visible Speech
IDENTIFIERS Clarke School for the Deaf; Nasality; Velar Control

ABSTRACT

A three-year project developed a computer-based system of speech-training aids for the deaf, formulated speech-training procedures that utilize the system, and evaluated these procedures. The first year was devoted to development of the system; in succeeding years it was employed in a remedial speech-training program at Clarke School for the Deaf. The system and its use at Clarke School are described, along with results of an attempt to evaluate its effectiveness. (Author/SK)

ED 100357

COMPUTERIZED SPEECH-TRAINING AIDS FOR THE DEAF

FINAL REPORT

Contract No. OEC-0-71-4670(615)

BBN Job No. 11580

Principal Investigators:
Raymond S. Nickerson
and
Kenneth N. Stevens

14 September 1974

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

IR 001 455

Contract Monitor:
Elwood Bland

Media Services and Captioned Films Branch
Division of Educational Services
Bureau of Education for the Handicapped

FINAL REPORT

Project Title: Development of Computerized Speech-Training Aids for the Deaf

Contractor: Bolt Beranek and Newman Inc.

Contract No.: OEC-0-71-4670(615)

Effective Date of Contract: 15 June 1971

Contract Expiration Date: 14 September 1974

Reporting Period: 15 June 1971 - 14 September 1974

Principal Investigators: Raymond S. Nickerson and Kenneth N. Stevens

Table of Contents

Page

Abstract..... iii

1. INTRODUCTION..... 1

2. CONTENT AND ORGANIZATION OF THE REPORT..... 2

3. SYSTEM DESCRIPTION..... 6

4. USER DOCUMENTATION..... 9

5. USE OF SYSTEM AT THE CLARKE SCHOOL..... 10
   - 5.1 Selection of Students..... 10
   - 5.2 Characteristics of Students..... 11
   - 5.3 Diagnosis..... 13
   - 5.4 Tutorial Instruction..... 19
   - 5.5 Unsupervised Drill..... 21
   - 5.6 Normative Data Collection..... 22

6. EVALUATION OF EFFECTIVENESS OF SPEECH TRAINING USING THE SYSTEM OF DISPLAYS..... 23
   - 6.1 Training Session Documentation..... 23
   - 6.2 Subjective Teacher Assessment of Overall Progress of Each Student..... 29
   - 6.3 Objective Data on Relevant Attributes of Test Utterances..... 31
   - 6.4 Subjective Judgments of Test Utterances..... 33
   - 6.5 Measures of Intelligibility..... 34

7. CONCLUSIONS..... 35

8. RECOMMENDATIONS..... 40

9. ACKNOWLEDGMENTS..... 42

Abstract

This is the final report of a three-year project to develop a computer-based system of speech-training aids for the deaf, to formulate speech-training procedures that utilize the system, and to evaluate these procedures. The first year of the project was devoted to the development of the system. During the second and third years the system was used in a remedial speech-training program at the Clarke School for the Deaf. This report, along with several others to which it refers, describes the system, its use at the Clarke School, and the results of an attempt to evaluate the effectiveness of the approach to speech training that it represents.

1. INTRODUCTION

This is the final report of a project to develop and evaluate a system of computer-based speech-training aids for the deaf under Contract No. OEC-01-71-4670. The project was initiated on 15 June 1971 and completed on 14 September 1974.

To our knowledge, this project represents the first attempt to use a computer to generate real-time visual displays of speech properties for the purpose of facilitating the teaching of speech to deaf children. From the beginning of the project we have felt that the best way to proceed would be to develop a prototype system that could be easily modified, to get it quickly into the hands of users, and then to develop the system's capabilities further in accordance with the insights gained as a result of trying to use it. Accordingly, the basic system was configured and the initial programming done during the first year of the project, and then the system was installed at the Clarke School for the Deaf where it has been used on an almost-daily basis for two years. The system has continued to evolve, and many of the characteristics of the displays that are currently operational are the results of suggestions made by the users.

A particularly gratifying aspect of the project has been the smoothness of the interaction between the researchers and the teachers involved. Frequent and regularly scheduled meetings between the BBN and the Clarke School members of the project team have facilitated the exchange of ideas and assured a continuing commonality of purpose. Many writers have pleaded for closer collaborations between researchers and teachers on the problem of developing procedures for teaching speech to the deaf. Our experience convinces us of the reasonableness of this plea. Each group has much to learn from the other, and the problem is certainly difficult enough to require the best efforts of both.

2. CONTENT AND ORGANIZATION OF THE REPORT

Much of the work that has been done under this project has been documented in a series of progress reports and technical papers. The progress reports have been issued and submitted to BEH on a quarterly basis. These reports contain many details of the design and use of the system, and provide a chronological record of the course of the project over the three years of its life.

The technical papers that have been prepared are listed in Table 1. Most of these papers were prepared for the purpose of eventual publication. At the time of this writing, three of these papers have been published either in a professional journal or in a conference proceedings, or both; five have been submitted to journals for consideration for publication; and three more will be submitted shortly.

In addition to progress reports and technical papers, three other documents have been prepared. These are: (a) a set of narrative reports on the activities and achievements of individual students, prepared by the teachers, (b) an operators reference manual and (c) a draft curriculum manual. These documents are listed as numbers 12, 13 and 14 in Table 1.

Each of the documents listed in Table 1 should be considered as part of the final report on this project. A complete set of these documents is being forwarded to BEH; copies of individual reports are available on request. There were two reasons for preparing a number of separate documents rather than attempting to integrate everything in one report: (a) to produce some manuscripts that would be appropriate for submission to technical journals, and (b) to facilitate selective distribution of the results of this project to other researchers and/or prospective users of this, or a similar system.

Table 1. List of documents prepared under Contract No. OEC-01-71-4670

1. Nickerson, R. S. & Stevens, K. N. An experimental computer-based system of speech training aids for the deaf. Proceedings, Conference on Speech Communication and Processing, IEEE; and Air Force Cambridge Research Laboratories, April, 1972, 238-241.

2. Nickerson, R. S. & Stevens, K. N. Teaching speech to the deaf: Can a computer help? Proceedings, Association for Computing Machinery, August, 1972, 240-252; and IEEE Transactions on Audio and Electroacoustics, 1973, AU-21, 445-455.

3. Nickerson, R. S., Kalikow, D. N., & Stevens, K. N. A computer-based system of speech-training aids for the deaf: A progress report. BBN Report No. 2901, September, 1974. Abbreviated version published in AFIPS Conference Proceedings, 1974, 43, 125-126. Also submitted to American Annals of the Deaf.

4. Stevens, K. N., Kalikow, D. N., & Willemain, T. R. The use of a miniature accelerometer for detecting glottal waveforms and nasality. BBN Report No. 2907, September, 1974. Accepted for publication, Journal of Speech and Hearing Research.

5. Stevens, K. N., Nickerson, R. S., Boothroyd, A., & Rollins, A. Assessment of nasality in the speech of deaf children. BBN Report No. 2902, September, 1974.

6. Nickerson, R. S., Stevens, K. N., Boothroyd, A., & Rollins, A. Some observations on timing in the speech of deaf and hearing speakers. BBN Report No. 2905, September, 1974.

7. Stevens, K. N., Nickerson, R. S., Rollins, A., & Boothroyd, A. Use of a visual display of nasalization to facilitate training of velar control for deaf speakers. BBN Report No. 2899, September, 1974.

8. Boothroyd, A., Nickerson, R. S., & Stevens, K. N. Temporal patterns in the speech of the deaf--An experiment in remedial training. Report from the Research Department, Clarke School for the Deaf S.A.R.P. #15, September, 1974.

9. Boothroyd, A., Archambault, P., Adams, R. E., & Storm, R. D. Use of a computer-based system of speech analysis and display in a remedial speech program for deaf children. (Based on a panel presentation to the Massachusetts Speech and Hearing Association, May, 1974); Clarke School for the Deaf S.A.R.P. #14, September, 1974.

Table 1 (continued)

10. Boothroyd, A. Teacher/researcher interaction--A model and an example. Proceedings of the 46th Meeting of the American Instructors of the Deaf, 1974, 122-129. U.S. Printing Office, Washington, D.C.

11. Nickerson, R. S. Speech training and speech reception aids for the deaf: A review. To appear as a BBN report.

12. Adams, R. E. & Storm, R. D. Narrative reports on students who received speech training with the BBN system. Report from the Research Department, Clarke School for the Deaf S.A.R.P. #16, September, 1974.

13. Rollins, A. M., Kalikow, D. N., & Nickerson, R. S. Operators Reference Manual and Software Documentation. BBN Report No. 2904, September, 1974.

14. Nickerson, R. S., Stevens, K. N., Boothroyd, A., Adams, R. E., & Storm, R. D. Draft curriculum manual: Review of speech problems of the deaf and preliminary guide to the use of the BBN system of speech training aids. BBN Report No. 2911, September, 1974.

The body of the final report proper consists primarily of overview and summary-type information, and "pointers" to the other documents where more detailed information regarding specific aspects of the project can be found. The final section contains a number of conclusions that we have drawn from the experience and results that have been obtained over the three years of the project, and makes several recommendations concerning directions that might be taken by future efforts to develop speech-training aids for the deaf.

3. SYSTEM DESCRIPTION

The system has been described in detail in papers numbered 1, 2, and 3 in Table 1. Only a brief summary description will be given here.

As illustrated in Fig. 1, speech information is acquired via a voice microphone and a miniature (BBN Model 501) accelerometer. The accelerometer is attached to the nose to acquire information concerning nasalization, and to the throat to obtain voice fundamental frequency. The output from the microphone is fed into a bank of 19 band-pass filters covering the frequency range from 100 to 6560 Hz. When the accelerometer is attached to the nose its output is fed into a "nasalization" detection circuit; when it is attached to the throat its output is fed into a pitch extractor.

The output of each of these "preprocessing" components--the nasalization circuit, the pitch extractor, and each of the 19 filters--is sampled by the computer (a Digital Equipment Corporation PDP/8E) every 10 milliseconds. These samples are continuously stored in an area of the computer's memory that we refer to as a "ring buffer." The size of the ring buffer is limited by the amount of computer memory that is not required for other purposes. Typically, we have used a buffer large enough to hold a representation of a speech sample of two seconds' duration. The data that are acquired at any given instant replace the oldest data in the buffer, so that the system always contains a record of the speech produced during the immediately preceding two seconds. By pressing a button on the control panel, the user can, at any time, capture the speech that is currently represented in the buffer. When the capture button is pressed, the recording of new speech data stops until another button is pressed to indicate that it should be resumed.

BEST COPY AVAILABLE

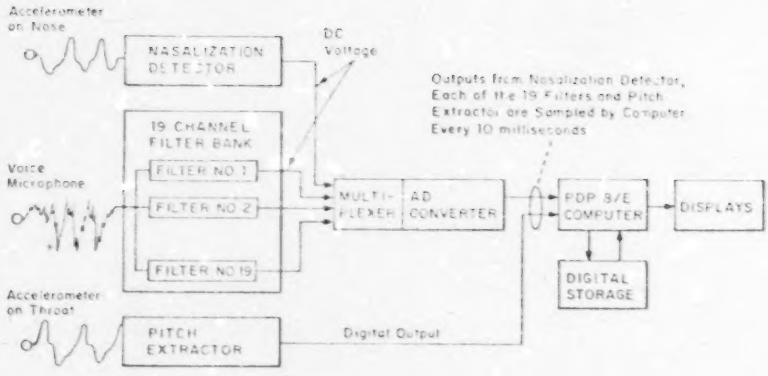


Fig. 1. Simplified block diagram of the system.

The data that are stored in the ring buffer provide the basis for a variety of displays. These displays can be operated in real time, providing the user with a visual representation of specified aspects of his speech (e.g., voicing, amplitude, fundamental frequency, nasalization) as he speaks. They also can be frozen and/or replayed for inspection after the fact.

Four different display programs have been developed. One presents the child with a game-like situation, permitting him to "score baskets" by successfully accomplishing certain vocal exercises. A second program represents certain speech parameters in the features of a cartoon face. A third provides information about the spectral properties of sounds by means of forms whose shapes depend upon the way the acoustic energy is distributed over the voice-frequency range. A fourth program provides the user with the ability to display various speech parameters as functions of time.

In addition to the display programs, a fifth program has been developed for the purpose of facilitating the analysis of recorded speech samples. This program permits the user to listen to a speech recording, to capture a segment of interest, to display certain parameters of the captured segment, and to obtain various objective measurements from the program on demand.

For more details concerning the system the reader is referred to the reports mentioned above.

4. USER DOCUMENTATION

The user documentation that has been prepared is of two types: (a) instructions concerning the mechanics of loading programs into the computer, setting program parameters, modifying programs, and so on, and (b) guidance and suggestions concerning the use of the displays in training sessions. This documentation is contained in two manuals: an Operators Reference Manual and a Draft Curriculum Manual (reports numbered 13 and 14, respectively, in Table 1).

The tentative nature of both documents, particularly the latter, is recognized. We are keenly aware of the fact that the guidance and suggestions that we have to offer concerning training are based on far too little experience to warrant being considered anything but tentative. The curriculum manual, therefore, is identified as a "draft" in order to emphasize the fact that we view it as, hopefully, something upon which to build. We think that much has been learned as a result of using this system in speech training classes over the past two years. However, anticipating the final section of this report, we may note that one of the main conclusions that we have drawn from that experience is that current speech processing and display-generation technology hold promise for the problem of teaching speech to the deaf, but a great deal more effort will be required by both researchers and teachers in order to learn to exploit this technology fully. Many of the most urgent and difficult problems appear to be less technical than pedagogical.

5. USE OF SYSTEM AT THE CLARKE SCHOOL

During the course of this work, 42 students (24 in 1972-73, 18 in 1973-74) at the Clarke School were given regular speech training using the system of displays. For most of these students, this training was primarily on a one-to-one basis with a teacher. During the final months of the 1973-74 school year, six of the 42 students also participated in unsupervised drill with the system.

5.1 Selection of Students

The selection of students for participation in this program was made on the basis of several considerations:

(a) The types of speech problems of the students. It was desired to work with several diverse types of speech problems, so that experience with these problems could be obtained and so that the effectiveness of training with the system could be assessed for different types of displays. Furthermore, some students had severe problems with a number of different aspects of their speech, whereas others had relatively minor difficulties with one or two speech attributes. In all cases, the speech problems of the students were such that it was expected that work with the system could be beneficial to their speech. Among those selected for training with the system were several students who had specific, obvious and severe speech problems which had not responded significantly to conventional training methods.

(b) Age and academic level of the students. The ages of the participating students ranged from 8 to 17. Work on an experimental basis was carried out for a short period of time during the summer of 1973 with two younger children, ages 4 and 6. Except for these two children, an effective lower age limit on participants was

imposed by the Lower School policy of not attempting one-to-one analytic speech work with children in their first and second years of school.

(c) Hearing level. Of the 42 students in the age range 8-17, one was severely deaf, 8 were profoundly deaf, but with high-frequency residual hearing, 17 were profoundly deaf with only low-frequency residual hearing, and 17 were possibly totally deaf. The definitions of these categories of deafness are given in Table 2.

(d) Availability of students for scheduling. Students were scheduled for tutoring with the system during the time of their regular speech class. Thus, scheduling problems imposed some limitations on the selection of students for this program.

(e) Individual requests by teachers and students. As the existence of the computer-based system and its capabilities became known in the school, occasional requests for tutoring with the system came from students and from teachers. Within the constraints described above, attempts were made to accommodate these requests.

5.2 Characteristics of Students

The students were approximately equally divided between girls (24) and boys (18). Most of the students (28) were tutored for a 7-week period, ten were in the program for 14 weeks, and only four were involved for over 14 weeks. There were somewhat more students in the 8- to 9-year-old group (15) and fewer in the 16- to 17-year-old group (4) than in the other age ranges.

Table 2. Definition of categories of hearing loss

severe: better ear three-frequency average (500, 1000, 2000 Hz) of 90 dB I.S.O. or less.

profound with high-frequency hearing: better ear three-frequency average in excess of 90 dB I.S.O.; thresholds at 2 or 4 kHz of 110 dB or less.

profound with low-frequency hearing: better ear three-frequency average in excess of 90 dB I.S.O.; thresholds at 2 and 4 kHz in excess of 110 dB, but thresholds at 125, 250, or 500 Hz less than 60, 75, or 100 dB, respectively.

profound, possibly total: better ear three-frequency average in excess of 90 dB I.S.O.; no measurable pure-tone responses, or responses are probably attributable to touch (i.e., in excess of 60, 75, and 100 dB at 125, 250, and 500 Hz, respectively).

Table 3 shows the numbers of students who worked on specific types of speech problems. Of the 42 students, 21 concentrated on only one problem, 16 worked on two problems, and five worked on three problems. The table shows that the problems receiving most attention were timing, intonation, and hypernasality.

5.3 Diagnosis

A crucial component of any speech training program is the diagnosis of the speech problems of particular students selected for training, and the formulation of training objectives for these students. During the first year of work with the system at the Clarke School, initial diagnosis was accomplished primarily on the basis of subjective assessment by the teachers--the classroom speech teachers, the speech coordinator of the school, and the research team involved in using the system to aid in speech training. Each student's speech problems were characterized in terms of several categories: timing problems; problems with pitch range and control of pitch; problems with velar control, leading to hyponasality, hypernasality, or both; problems with voice quality, particularly breathiness or harshness; and problems with articulation. Problems with articulation were usually specified in greater detail, such as difficulty with /s/, inadequate vowel range, inadequate diphthongization, problems with consonant clusters (or blends), etc.

Following the initial diagnosis, recommendations were made that a student receive training in one of these areas of emphasis. Initial work with the student in tutorial sessions with the displays usually confirmed the initial subjective diagnosis, and from this evaluation it was possible to formulate short-range objectives for the training sessions. Some of the students selected for tutorial work in the first year were those with very bad speech whose problems had long resisted remediation.

Table 3. This table shows the speech problems dealt with in one-to-one tutoring, and the number of students who worked on each type of problem.

BEST COPY AVAILABLE

Area of Tutorial Emphasis		Timing	Pitch		Voice Quality		Articulation		
		Rate and Rhythm	Average	Intonation	Larynx Function	Hyper-nasality	Nasals	Others*	Totals
Number of Students	Priority: First	16	1	10	2	11	1**	1	42
	Second	9	1	7	1	2	0	1	21
	Third	1	1	1	0	1	0	1	5
Totals		26	3	18	3	14	1	3	
		26	21		17		4		

* Vowels, diphthongs, and /s/.

** This student learned to articulate /n/.

During the second year of work with the system at the school, diagnosis of speech problems and formulation of training objectives were carried out on a more systematic and quantitative basis. Each student who was a potential candidate for work with the system was given an initial evaluation based on objective measurements and on subjective assessments for a specified body of speech material generated by the student. Some of the data were collected on-line by making measurements on the displays with the student present; other measurements and judgments were made subsequent to the initial recording session. This evaluation consisted of the following:

(a) Measurements of average fundamental frequency (F_0) and F_0 range for a selected sentence.

(b) Measurements of HL reading (see Report #3 of Table 1). A high reading on this scale is a tentative indication of breathy voice quality.

(c) Comparison of "nasalization readings" for nonnasal vowels and nasals in selected one- and two-syllable words.

(d) Measurements of durations of selected speech events in a simple paragraph read by the student.

(e) Subjective judgments by the two research teachers of various speech attributes relating to voice quality, pitch, and timing, based on listening to the recorded paragraph.

(f) Subjective judgments of the articulation of selected vowels and consonants occurring in a list of words recorded by the student.

On the basis of all of these data, a "dossier," describing the speech of each student, was prepared. This included also

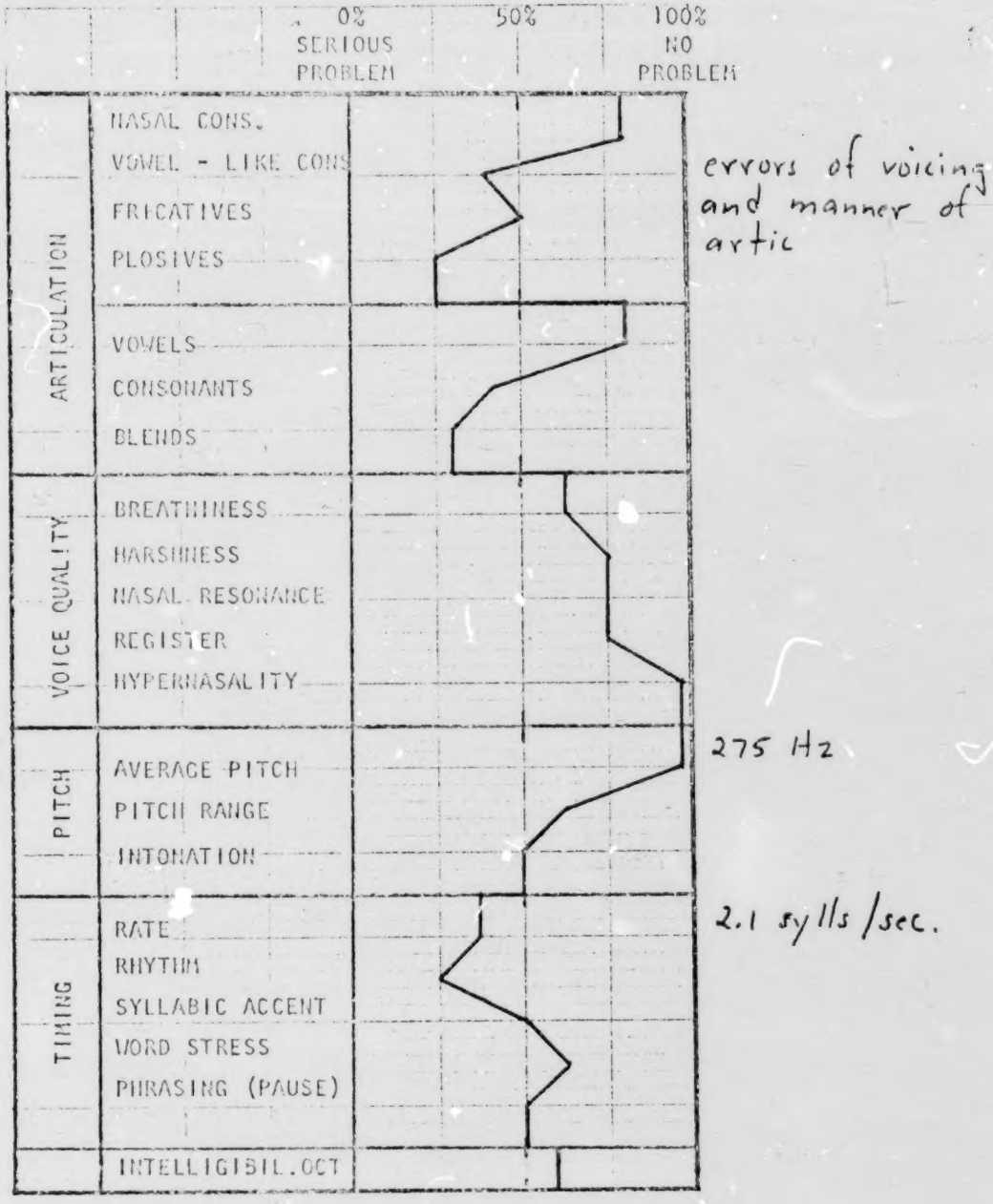
information on the student's audiogram and intelligibility scores achieved in the regular intelligibility testing program at the school. One of the items in this dossier was the "speech profile," shown in Fig. 2. Problems with articulation are assessed by critical listening to the items in the word list; problems with voice quality, pitch, and timing are determined by rating of the various attributes indicated at the left of the figure on a 5-point scale. The percent intelligibility (from the standard Clarke School tests) is noted at the bottom of the profile. Notations concerning average fundamental frequency, range of fundamental frequency and syllabic rate during the sentences are made to the right of the chart.

The data in Fig. 2 and the objective measurements listed above were supplemented, in the case of some students who were diagnosed to have specific speech problems, by additional objective measurements of timing, nasality, and HL value.

This material was used to provide a preliminary diagnosis of the speech problems of students selected for inclusion in the speech-training program during the second year at Clarke School, and to provide a basis for the establishment of initial training objectives. In some cases, the early training sessions revealed more detailed speech problems not uncovered in the initial material used for diagnosis.

In addition to being used to help diagnose the speech of the children who were to participate in the project, the system was occasionally used, at the request of a teacher, to provide a quantitative estimate of some property of the speech of a child (e.g., the fundamental frequency, the degree of nasalization, etc.) who was not participating in the program.

COMPUTERIZED SPEECH TRAINING AID PROJECT



SPEECH PROFILE

Articulation and suprasegmental features--subjective evaluation.

Fig. 2. Method of summarizing data from subjective evaluation of diagnostic speech sample. Written notations at right include data from objective measurements of fundamental frequency and timing.

Over the course of this project we have become increasingly aware of the importance of diagnosis in the overall speech-training program. A detailed diagnosis of an individual child's speech deficiencies is a prerequisite to the establishment of reasonable training objectives, and of training procedures that are likely to realize those objectives. It is also essential to the accurate assessments of progress that a child may be making as a result of training. We have expressed our views on this subject in Section 3 of the Draft Curriculum Manual.

It should be emphasized that the diagnostic procedure that is described above is only a tentative one, and in need of considerable refinement. It represents a step, however, in the direction of producing the type of quantitative information that will be necessary, we think, for truly effective training procedures. We believe that the development of a comprehensive, quantitative, easily-administered diagnostic procedure could represent a major advance on the problem of speech training for the deaf. If such a procedure could be developed, one can imagine that it might become as important a tool for education of the deaf as audiometric testing procedures currently are. A speech-proficiency profile could, in principle, represent for speech what the audiogram represents for hearing.

A second and extremely important potential benefit that would result from the widespread use of a standard speech-diagnosis procedure would be the accumulation over time of a comprehensive body of data that could be used to describe, in objective and quantitative terms, the characteristics of the speech of the deaf.

5.4 Tutorial Instruction

The instruction of the students in this program consisted primarily of tutoring on a one-to-one basis with both student and teacher in front of the displays. In the second year of the program, a few students were involved in unsupervised drill. The "self-instruction" mode will be described in the following section of the report.

Almost all of the tutorial instruction was carried out by two teachers. These individuals (Mr. Robb Adams and Mr. Robert Storm) were qualified teachers of the deaf at the Clarke School, and each was assigned half-time to this project. The remainder of their time was spent in classroom teaching. Before tutorial sessions with the system began at the school, these teachers received training in the use of the system and in some principles of acoustics and phonetics relevant to proper interpretation of the displays.

Some tutoring with the system was also carried out by the Director of Research at the school (Dr. Arthur Boothroyd), and, on occasion, by student teachers.

The training sessions consisted of various types of exercises that were graded in difficulty. The general approach was to work with the student until he could achieve a reasonable level of performance in one type of exercise before proceeding to the next. The teachers used some judgment in selecting tasks for students, however, in the interest of maintaining motivation in the students and avoiding long sessions in which they failed to achieve successes. (Training procedures are discussed in greater detail in documents 7, 8, 9 and 14 in the list of Table 1.)

The types of exercises prepared for the students depended, of course, on the aspect of speech that was being considered. For the most part, however, these exercises were of two general types:

(a) Work with "vocal gymnastics," in which the sounds produced by the student are nonspeech activities such as sustained vowels or consonants, simple nonsense syllables, or sequences of such syllables.

(b) Production of meaningful utterances of various lengths, selected to contain the speech gesture that is being worked on.

For the vocal gymnastics activities, it was possible in some situations to set objective criteria that the student was asked to achieve, such as producing an isolated vowel with nasalization reading below a specified value, or producing a sustained vowel with pitch in a specified range, or with a particular terminal fall in pitch. In other cases, the teacher generated a "target" pattern, and the student was required to match this pattern, the judgment of adequacy of match sometimes being a subjective one. Examples are the timing of a sequence of two nonsense syllables such as pa pá, or the generation of /s/ with the proper vertical spectrum pattern. Suggested sequences of training exercises for specific classes of speech problems are given in the Draft Curriculum Manual.

Initial orientation and training were, of course, carried out with the displays being generated in real time, yielding an instantaneous record of the student's speech activities. After this initial training, however, some exercises were carried out with the system in "delay mode," such that the student did not observe the display as he was producing the utterance. In this case, the display was put on the screen only after the student had had an opportunity to assess the adequacy of his utterance. This component

of the training activity was carried out in an attempt to "wean" the student away from reliance on the display and to accelerate the process of internalization of the student's skills and carry-over to spontaneous speech. This internalization process was emphasized only during the second year of use of the system at the Clarke School.

5.5 Unsupervised Drill

Although we did not originally plan for this particular system to be used by students without a teacher present, it became apparent during the course of the project that many of the students were quite capable of using the system in this way. Therefore, toward the latter part of this year we began experimenting with this possibility.

Six children, between the ages of 12 and 17, have used the system in this way. Each had participated in the tutoring sessions before using the system in "self-instruction" mode. The average numbers of tutorial and unsupervised drill sessions for these children were 43 and 36, respectively. In some cases, the two types of sessions were interspersed.

During a self-instructional session the child is on his own. He comes to the "orange room" (the name given to the room where instruction with the computer displays is given), turns on the computer, puts on the microphone and accelerometer, selects the display with which he is to work, practices exercises that he and a teacher have agreed upon, and then shuts the computer off and leaves, when he is done. Simple self-scoring procedures have been developed and are used by the children during these sessions to record their progress as they work toward specific objectives. The children seem to enjoy working with the system on their own; there certainly has been no evidence that they are intimidated by the fact that they are interacting with a computer.

Using the system in a self-instructional mode has several advantages, perhaps the most obvious of which is the fact that it does not require the expenditure of the teacher's time. It also has a danger, however. Because a teacher is not present and monitoring the performance of the child, there is the possibility that while he is working on one aspect of speech, other aspects, to which he is not attending, may be affected negatively. It is important, therefore, at least until adequate self-instructional procedures are developed, that the effects of self-instruction be checked periodically by teachers in order to insure that bad speech habits are not being practiced inadvertently. We do feel, however, on the basis of our experience to date, that self-instructional use of such systems holds very real promise. We suspect that for many children the optimal training strategy will involve some mixture of tutorial and self-instructional training.

5.6 Normative Data Collection

A final purpose for which the system has been used is to collect data that could be helpful in establishing quantitative training objectives with respect to certain aspects of speech. If one is to attempt to improve the speech of a deaf child with respect to some particular property or properties, it is necessary to know not only how the speech is deficient but what must be done to make it correct. Unfortunately, it often is not possible to determine from existing data exactly what one should work toward in the case of many properties of speech. We have, therefore, in some instances, collected data from normally-hearing speakers to help provide training objectives with respect to some speech properties. In particular, we have done this with respect to timing and nasalization. The results of these efforts are described in papers numbered 5 and 6 in Table 1. The use of these data in the establishment of training objectives and for the purpose of evaluating effects of training has also been documented in papers numbered 7 and 8 in Table 1.

6. EVALUATION OF EFFECTIVENESS OF SPEECH TRAINING USING THE SYSTEM OF DISPLAYS

The effectiveness of the speech-training program was assessed in a number of different ways. These include:

(a) Daily and weekly records maintained by the teachers of the objectives and accomplishments of the children for the individual training sessions.

(b) Subjective evaluation, by each teacher, of the overall progress of his students over the course of their entire training programs.

(c) Objective data on relevant acoustic characteristics of test utterances recorded before, during, and after, training.

(d) Subjective judgments on the utterances noted in (c), including judgments with respect to particular attributes of the speech (timing, nasality, etc.).

(e) Measures of speech intelligibility.

6.1 Training Session Documentation

As we have already noted, the speech of each child who participated in this project was diagnosed before training was begun. On the basis of this diagnosis the long-range goals for the child were established. These goals were relatively general, e.g., to improve the speech with respect to some particular aspect or aspects.

Specific daily objectives, consistent with the long-range goals, were set for each child on a week-by-week basis. These objectives were of the sort:

- Using short sentences,to articulate unaccented syllables at or slightly below his habitual pitch, while making accented syllables at least 50 Hz above his habitual pitch (for a child with a habitual pitch of 280 Hz), with the terminal fall being the lowest pitch for the utterance.

- To produce utterances with nonnasal consonants with a nasalization reading of under 50, and nasal consonants with a nasalization reading greater than 200.

- To produce two- and three-syllable utterances with different stress patterns, making the duration of the stressed syllable three to four times the duration of the unstressed syllable when the stressed syllable appears in final position, and one to two times as long when it appears in other than final position.

In order to document the progress of each child vis-a-vis these objectives, the teachers maintained daily records in which they described the exercises that were used during each training session, and indicated the degree of success realized by the child during that session. On occasion, these data were recorded systematically in such a way that an assessment could be made of the percent of trials that were successful, the number of trials needed before some criterion performance was achieved (such as five successive correct attempts), or the percent of trials in which the student was able to assess correctly the adequacy of his production (with the display in delay mode). During the latter part of the program, occasional snapshots of the display showing

the student's performance with particular utterances were made to supplement the written records.

A weekly summary of the information contained in the daily records was also prepared by each teacher for each of the children with which he was working. These summaries were reviewed and discussed by the project team during the meetings that were regularly held for the purpose of assessing the overall progress of the project and to revise training objectives for the individual students as progress warranted. An example of these summaries is shown in Fig. 3.

Both the daily records and the weekly summaries contain many data that are germane to the question of the effectiveness of this approach to training. This documentation demonstrates clearly that all of the students made some progress in the training sessions, in the sense that they were able to accomplish speech exercises or produce utterances for which they were not successful at the onset of training. Examples of data that show this type of progress are shown in Figs. 4 and 5. Figure 4, for example, demonstrates that a student who was initially unable to produce a sentence without nasalization was able after some training to correct this problem and consistently to generate the sentence with a raised velum. Figure 5 gives the pitch contour of a particular utterance with an inappropriate terminal rise before training and with a well-executed terminal fall after a period of training.

At the end of a student's training period (7 or 14 weeks for most students), a final summary was prepared for that student. The daily and weekly summaries were used as input material for this purpose. These summaries comprise the report listed as number 12 in Table 1.

Student's Name A.M. Tutor Robb Adams

Date(3) March 8-14, 1974 Week # 10 (5 sessions)

Emphasis: velar control and timing

Objectives:

To produce nonsense syllables and words with appropriate timing features. (Ratio of duration of unstressed syllables to duration of stressed syllables is approximately 1/3 to 1/4. Intraphrase gap is approximately 0 to 1/3 times the unstressed syllable.)

To produce utterances (nonnasal) under a nasality level of 50.

To produce utterances with appropriate timing and nasality features.

Progress:

A.M. produced utterances with appropriate timing and nasality features. Specific results are attached.

The student is inconsistent with his velar control. One day he will succeed with most utterances attempted and the next day he will have difficulty with most utterances attempted. Specific results are attached.

Timing and velar control were evaluated with the VN display in flow mode. Only a few utterances were attempted. Results are attached.

Results and Comments:

A.M. is more successful with nonsense syllables than real words because of the articulation requirements involved. It is a slow process working from nonsense to meaningful utterances that he can succeed with.

Success with velar control is sporadic, even with nonsense material. Certain sounds are usually nasal for the student (/oo/, /ee, /a-e/, /o-e/. /i-e/. /1/, /u-e/, /d/).

Because of limited exposure it is difficult to say at this time how effective the VN display will be for evaluating timing as well as nasality. The student needs to work until it is easy for him to articulate one-, two-, and maybe three-syllable utterances initially. Timing alone is not very difficult for A.M., but as the length or rate of an utterance increases, velar control and articulation fall apart.

Subsequent Activities:

Produce utterances with appropriate* nasality and timing features.

*Description of appropriate timing and nasality features appears under Objectives above.

Fig. 3. Weekly summary sheet - computer/speech program

1. Timing: nonsense syllables, phrases
   Criteria: ratio - duration of unstressed syllable to stressed syllable = 1/3 to 1/4.
   gap (intraphrase) - approx. 0 - 1/3 times unstressed syllable.

pu páCr)	±|10
páCr) pu páCr)	±|2 |4/1 |1 |1 |2|1 |10
pu pu páCr)	±|10
pu páCr) pu	±|10
A M	±|2 |1 |8 |1 |5

2. Velar control: nonsense syllables, words, phrases
   Criteria: to produce nonnasal utterances below a nasality level of 50.

pu páCr)	±|10 , ±|2 |4 |1 |2 |5 |1 |7
páCr) pu	±|5
pu pu páCr)	±|20
paCr)	±|10
Hello.	±|20-30
How are you?	±|1/2 |1 |1 |1 (criterion - under 100)
Goodbye, A	±|1 |6 |2 |2 |3 |10 |2
óo	±|2 |1 |6
óot	±|2 |1 |3 |1 |10

Sheet 2 of Fig. 3.

BEST COPY AVAILABLE

3. Velar control and timing
Criteria – same as sections #1 and #2.

boot	$\frac{+}{0}\|h\|_1\|^6\|_1\|^6\|_1\|^{13}\|_1\|^2$
my boot	$\frac{+}{0}\|^4\|_1\|^4\|_2\|^3\|_1\|^{19}$
goodbye	$\frac{+}{0}\|^{10}$

Sheet 3 of Fig. 3.

BEST COPY AVAILABLE

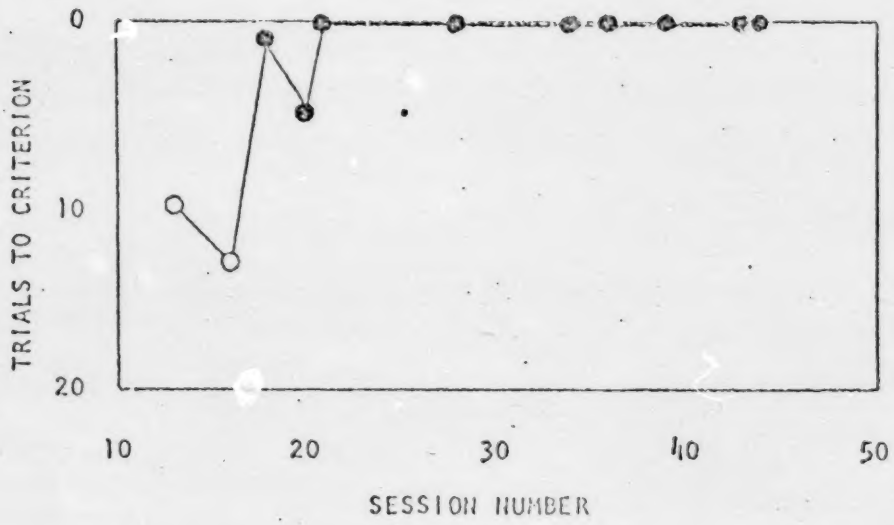


Fig. 4. Performance of a deaf student in producing the sentence "I live at Clarke School" with nasalization reading (output of accelerometer attached to nose) below a specified value. The ordinate is the number of trials necessary before the first of 5 consecutive successes was achieved.

BEST COPY AVAILABLE

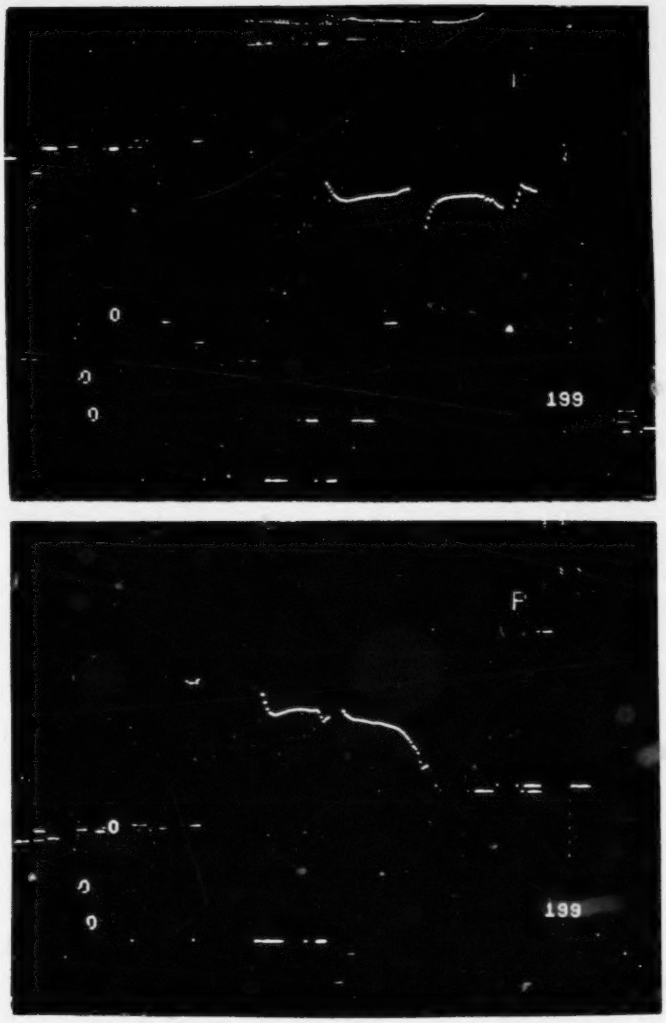


Fig. 5. The photograph at the top shows the presence of a sudden pitch rise at the termination of the utterance "Audrey." This inappropriate behavior was characteristic of this student's speech. The bottom picture shows that, after some training, the student eliminated the pitch jump and produced a natural terminal fall in pitch.

6.2 Subjective Teacher Assessment of Overall Progress of Each Student

Following completion of each student's participation in the project, his performance was rated by his tutor with respect to four aspects:

a. improvement of vocal gymnastic skills
b. incorporation of these skills into rehearsed speech
c. carryover to unrehearsed speech (formally presented)
d. carryover to spontaneous speech

Ratings were made on a 5-point scale as follows:

0. no improvement noted
1. a small improvement noted, but of questionable significance
2. a moderate and significant improvement noted
3. considerable improvement noted, but student failed to meet all criteria or showed inconsistent performance
4. student showed considerable improvement and learned to meet criteria consistently and with ease

Ratings were assigned on the basis of the goals that were set for tutoring. Thus, two students who received the same rating were not necessarily accomplishing tasks of the same difficulty.

The results of the ratings have been discussed in the paper numbered 9 in Table 1. Table 4 shows the ratings of the students on each of the four dimensions for the speech area of primary emphasis. A comparison of ratings on the four dimensions shows that the dimensions of vocal gymnastics and rehearsed speech were rated significantly higher than the two dimensions related to carryover. This is in keeping with our observations that the learning of gymnastic skills and their incorporation into rehearsed speech could generally be accomplished

BEST COPY AVAILABLE

Table 4. Distribution of tutors' subjective ratings of progress for 42 students on four rating dimensions. Ratings on dimensions C and D (both indicants of carryover) were significantly poorer than those on dimensions A and B. (Differences significant at the 1% level, based on X^2 approximation to the Kolmogorov-Smirnov test, 2 d.f.)

Rating dimension	Tutor ratings					
	No progress ⟷ Considerable progress					Mean Rating
	0	1	2	3	4	
A - Vocal gymnastics	0	3	11	16	12	2.9
B - Rehearsed speech	1	5	16	13	7	2.5
C - Unrehearsed speech	2	18	13	7	2	1.7
D - Spontaneous speech	9	16	11	5	1	1.4
Totals	12	42	51	41	22	2.1
	Number of students					

with little difficulty in a 7-week period. The generalization of these skills to unrehearsed and spontaneous speech required much longer periods of tutoring and even then was accomplished by only a small number of students.

Further analysis of the data revealed tendencies for girls to be rated higher than boys, for older students to be rated higher than younger students, and for students with more residual hearing to be rated higher than students with less residual hearing. These were only trends, however. None of the differences reached the 5% level of significance.

Significant differences were demonstrated between the ratings for students participating in Years 1 and 2 of this project, with those on Year 2 scoring higher. It will be recalled that the students in Year 1 tended to be selected from a pool of "difficult" cases, whereas a more representative sample was chosen for Year 2. It should also be pointed out that much more attention was given to internalization and carryover in Year 2 and it was only on the dimension of carryover to spontaneous speech that the difference between Year 1 and Year 2 ratings reached the 5% level of significance. A further factor or relevance to this observation is the increased experience of the tutors during Year 2 of the project.

A comparison of the students on the basis of speech problems revealed few differences between pitch, nasality and timing groups. The pitch group was, however, rated significantly lower on the dimension of carryover to spontaneous speech (difference significant at the 5% level).

6.3 Objective Data on Relevant Attributes of Test Utterances

Before training was commenced, a series of utterances was recorded for each student involved in the training program. The

recordings were made on two channels--one channel was the microphone signal, and the other was the signal from the accelerometer attached to the nose. The recorded material consisted of: (a) a list of 33 monosyllabic words, each elicited from the students by showing a picture; (b) a series of short phrases and sentences; (c) an approximation to spontaneous speech obtained by having the student tell the story described by a sequence of pictures; and (d) a list of six 10-word sentences drawn from a larger list that is used routinely to evaluate the intelligibility of the speech of students at the Clarke School. The same set of material was recorded within a few days after speech training with the system was terminated. These recordings were available for making objective measurements of various attributes of the utterances of the students. The recordings could also be used to obtain subjective judgments of different aspects of the speech, and to assess the intelligibility of the speech before and after training.

The pre- and post-training recordings just described were occasionally supplemented by recordings of special speech material that was designed to assess the progress of students with regard to particular speech attributes. For example, the reading of a short paragraph was recorded for the purpose of assessing certain aspects of timing. Some of these recordings have been analyzed or evaluated for the purpose of specifying the properties of the speech of deaf students in comparison with normally-hearing individuals, in order to provide a rational basis for the diagnosis of the deaf students. The results of some of these analyses are presented in reports listed as numbers 5 and 6 in Table 1.

Comparisons between the before- and after-training recordings and the special speech samples recorded at various times during training were made for the purpose of assessing the effects of training on the speech of the deaf students. The results of some

of these comparisons are presented in reports listed in Table 1 as numbers 7, 8 and 9.

Detailed data of this kind were obtained from eight students who received training in timing, nine who were trained in velar control, and one whose tutorial sessions were concerned with pitch control. The data for timing showed that after some training students achieved significant reductions in the duration of unstressed syllables and within-phrase gaps, relative to the durations for the pre-training recording (report No. 8). Most students who received training with velar control were able to produce more nonnasal words with a raised velum after training than before--both words in isolation and words in sentences--although not all students showed consistent improvement for words in both types of utterances (report No. 7). The student for whom F_0 measurements were made demonstrated an ability to eliminate inadvertent pitch jumps at the ends of words and to produce an appropriate terminal fall in pitch (report No. 9). In summary, then, the available results demonstrate clearly that training was effective in modifying the speech of these students, making it more similar to that of hearing speakers with respect to certain objective measures of specific aspects.

6.4 Subjective Judgments of Test Utterances

In addition to being subjected to a variety of objective measurements, some of the before- and after-training speech samples and special samples produced at different times during training were also used in tests in which listeners were asked to evaluate utterances in terms of specific properties. In particular, ratings of timing and rhythm, and of velar control, were obtained. The results of these tests are presented in reports listed as numbers 7 and 8 in Table 1. Again, the results indicate improvements in

the speech of the participating students with respect to specific aspects, insofar as the quality of these aspects can be judged by listeners. The judgments of timing showed significant improvement, whereas the judgments of velar control, while demonstrating some improvement for some students, showed considerable variability between listeners.

6.5 Measures of Intelligibility

The ultimate goal of speech training is, of course, to increase the intelligibility and improve the quality of the student's speech. Whether it is reasonable to expect to make noticeable improvements in intelligibility as a result of 15 to 20 minutes of training a day over a period of a few weeks is questionable. Nevertheless, intelligibility measures were obtained for before- and after-training speech samples of several of the students who participated in this project. The results of these comparisons are given in reports numbered 7 and 8 in Table 1. The results were not consistent from student to student, and such gains in intelligibility as were found were small. The measure that was used--number of words correctly perceived--was relatively crude, however, and because of its sensitivity to numerous exogenous variables, would be a reliable indicant of changes only if the increases were large.

7. CONCLUSIONS

We present here the conclusions that we have drawn as a result of our involvement in this project. Some of these conclusions follow directly from the data that are presented in the various project reports. Others are more in the nature of opinions that have been shaped by the combined effects of several factors, e.g., extensive review of the literature, observations of speech-training activities (both with and without the system), conferences with teachers of the deaf, and interactions with deaf children.

a. The project has demonstrated the feasibility of a close interaction between researchers and teachers in the development of innovative speech-training aids and procedures for their use. We do not wish to suggest that this is the first time researchers and teachers have interacted effectively in this way. It does seem to be the case, however, that such collaborations are more often preached than practiced. Moreover, we have become increasingly convinced of the desirability, if not the necessity, of such an approach.

b. Students within the age range of those who participated in this study (8 to 17 years) have little, if any, difficulty in interpreting displays such as those that were used in this project. In particular, the differences between acceptable and unacceptable patterns were readily learned.

c. Some evidence was obtained (experience with two children, ages 4 and 6, during the summer of 1973) that some of these displays might be used to advantage in the teaching of much younger children.

d. Many teenage students are capable of using such a system by themselves in a self-instructional mode. Such self-instructional use probably should be attempted, however, only after the student has had some experience in the more conventional tutored situation. Moreover, when a student does use the system by himself, his performance should probably be monitored periodically by a teacher in order to assure that the practice that he is getting is having the desired effects. For many children an optimal training program may include a mix of tutored and self-instructional sessions.

e. The approach to speech training that was taken in this project appeared to have some motivational advantages. In general, the children worked willingly during the training sessions. Also, several requests were received from students who had participated at some time during the project to return for additional tutoring, or for use of the system for self instruction.

f. The displays and training procedures that were developed proved to be effective in helping the students to attain a variety of specific training objectives. These objectives had to do with improving speech with respect to certain specific features that had been diagnosed as being defective. The following examples illustrate the types of speech problems for which some degree of success was obtained:

Timing. With the help of displays of temporal patterns of speech, students were trained to produce phrases and sentences with improved timing, including a reduction in pauses, a shortening of the duration of unstressed vowels and an increase in syllable-production rate.

Velar control. With the help of a display of nasalization (based on the amplitude of vibration on the surface of the nose), students were trained to produce vowels with reduced nasalization and to gain more appropriate temporal control of the velum in words and phrases.

Pitch control. Training with the assistance of displays of fundamental-frequency contours was effective in increasing the ability of some students to speak with the pitch in a more appropriate range, to produce a fall in pitch at the end of an utterance, or to reduce the incidence of inadvertent pitch jumps.

Breathiness. A display that indicated excessive low-frequency energy for vowels, and, by inference, an inappropriately spread posture of the glottis, was used with two students to improve their ability to produce vowels without excessive breathiness.

Articulation. Displays of spectral distribution were used effectively with some students to improve the pronunciation of some consonants (in particular /s/ and /š/) and of certain vowels.

For all of these aspects of speech, improvement in speech skills was most marked for training exercises and utterances that were rehearsed, and was less (and sometimes quite small) for unrehearsed meterial and for spontaneous speech. There were large differences in the improvements achieved by different students.

Inasmuch as paired-comparison control group experimentation was not attempted, the data do not warrant the conclusion that improvements such as those that were obtained could not be realized as readily without the use of these displays. There is abundant evidence in the literature, however, that conventional

speech-training techniques have not been very successful. Therefore, evidence that a procedure works in the sense that it results in measurable improvements in speech can be taken as at least tentative evidence that the procedure is effective relative to others that have been used.

g. The teacher-users of the system learned a considerable amount about speech by virtue of being exposed daily to visual representations of various aspects of both normal and defective speech. In particular, much of value was learned concerning specific problems in the speech of the deaf and about the effectiveness of various approaches to training designed to alleviate those problems. We believe that such a system could play a very useful role in teacher-training programs.

h. The system proved to be helpful in the quantitative diagnosis of specific speech problems and in the objective assessment of progress that individual students were making with respect to these problems. Because of this more quantitative approach to the assessment of the speech of the students and to the evaluation of the progress of students, it was possible to be more precise than is typically the case in the specification of a graded sequence of exercises and training tasks.

i. There is a need, however, for the development of more thorough and more quantitative procedures for diagnosing the speech deficiencies of individual children. The criticalness of this need became increasingly apparent to us over the course of the project.

j. There is also a need for more complete normative data to provide the basis for the establishment of specific training objectives, given specific deficiencies,

k. A third need is for more intensive research aimed at the determination of the relationship between the objective and perceptual properties of speech, and especially the speech of the deaf. In particular, there is a critical need to know how the objectively measurable properties of speech relate to its intelligibility and perceived quality.

8. RECOMMENDATIONS

We believe that the philosophy that has guided this project--that of involving researchers and teachers in a collaborative effort to evolve an effective system of speech-training aids--has proven to be a sound one. The system that has resulted from this effort has been applied effectively to the speech-teaching task. And, although the system has continued to evolve over the course of the project, its potential, we believe, has not come close to being fully realized. Many ideas about further display possibilities have not been explored simply because of resource limitations. What has been learned, however, can provide a solid basis for the furtherance of this approach. Moreover, the system remains highly flexible and the programming can be readily modified so as to provide a vehicle for the testing of new ideas.

The next step that should be taken, we believe, is to involve a larger and more heterogeneous group of researchers, teachers, and deaf students in a coordinated effort to extend this approach. A wider exposure of the system to potential users is probably the best way not only to facilitate the further development of the system's capabilities but also to develop new methods that will make the best use of the capabilities that the system provides. Enough has been learned about the potential of the system and about procedures for using it to warrant the installation of systems similar to the one that has been used at Clarke at a small number (say, three or four) of other schools for the deaf, and to engage the users in a collaborative effort to extend this approach. We emphasize "collaborative" because we feel strongly that the establishment of a consortium with a common goal, common tools, and a continuing interchange of ideas and information is the way to maximize the chances of significant progress in this area.

Finally, there is a need for the development of training procedures that are designed to take advantage of the speech-processing and display-generation capabilities that modern technology provides. We have become increasingly convinced that the most serious impediments to the development and effective use of instrumental training aids are pedagogical uncertainties rather than technical limitations. There now exists a technology for processing speech and generating real-time displays that cannot be fully exploited until more is known about how speech should be taught.

9. ACKNOWLEDGMENTS

Many people have contributed significantly to this project. The BBN project team has changed from year to year as the needs of the project have changed. Each of the following individuals has made a substantial contribution during one or more phases of the effort: Juan Anguita, Paul Coughlin, Douglas Dodds, Susan Graesser, Charlene Long, Mary Mullarkey and Thomas Willemain. Daniel Kalikow and Ann Rollins have been major contributors to the project over its entire life.

The Clarke School project team was headed by Arthur Boothroyd and included Robb Adams, Patricia Archambault, Lawrence Prusak and Robert Storm. We are especially pleased to acknowledge the superb way in which the entire Clarke School community--administration, faculty and students--supported this effort.

Ideas, encouragement and constructive criticism have come from many quarters. We are particularly indebted to Lois Elliott, who served as contract monitor during the formative stages of the project, and who was the source of several helpful suggestions concerning training and evaluation procedures.

We also are most appreciative of the patient typing and retyping of the substantial amount of documentation associated with this project by Florence Maurer and Mary Tiernan at BBN and Jan Storm at the Clarke School.